

# Relatório Técnico: Análise de Sentimentos em Tweets sobre Clubes de Futebol Ingleses

## 1. Introdução

Este documento explica a implementação de um sistema de análise de sentimentos para avaliar a percepção pública sobre clubes de futebol ingleses, seguindo rigorosamente o roteiro especificado. O sistema utiliza Processamento de Linguagem Natural (PLN) com a biblioteca spaCy e a extensão SpacyTextBlob.

## 2. Estrutura do Código e Funcionalidades

### 2.1 Configuração Inicial (Passos 1-3)

```
!pip install -U spacy==3.8.0 spacytextblob==4.0.0
!python -m spacy download en_core_web_sm==3.8.0
```

- Finalidade: Instala versões específicas e compatíveis dos pacotes para evitar conflitos.

#### Mecanismo:

- spacy: Biblioteca principal para PLN
- spacytextblob: Extensão para análise de sentimentos
- en\_core\_web\_sm: Modelo de linguagem em inglês

### 2.2 Pipeline de Processamento (Passo 3)

```
nlp = spacy.load('en_core_web_sm')
nlp.add_pipe('spacytextblob')
```

#### Fluxo de Trabalho:

- Carrega o modelo de linguagem
- Adiciona o componente de análise de sentimentos ao pipeline

**Saída Esperada:** Lista de componentes (['tok2vec', 'tagger', ..., 'spacytextblob'])

### 2.3 Validação do Sistema (Passos 4-5)

```
doc = nlp("This is a wonderful campsite...")
{
  'score': 0.6, # Valor no intervalo [-1, 1]
```

```
'interpretation': 'Positivo'
}
```

#### **Lógica de Classificação:**

- score > 0.1 → Positivo
- score < -0.1 → Negativo
- Demais casos → Neutro

**Normalização:** Valores são garantidos no range [-1, 1] com  $\max(-1.0, \min(1.0, \text{valor}))$

## 2.4 Análise de Tweets (Passos 6-7)

for tweet in tweets:

```
doc = nlp(tweet)
{
    'tweet': texto_resumido,
    'score': -0.825,
    'sentiment': 'Negativo',
    'intensidade': 'Forte' # Baseado em |score| > 0.5
}
```

**Tratamento de Erros:** Try-except para cada tweet individual

**Saída Estruturada:** Inclui:

- Trecho do tweet (70 caracteres)
- Score numérico
- Classificação qualitativa
- Intensidade do sentimento

## 3. Fluxo de Processamento

1. Pré-processamento: Instalação e configuração do ambiente
2. Validação: Teste com texto controlado
3. Análise em Lote: Processamento dos tweets reais
4. Saída: Resultados formatados para interpretação

## 4. Dicionário de Campos de Saída

Campo	Tipo	Descrição	Exemplo
score	float	Valor entre -1 (negativo) e 1 (positivo)	0.725
sentiment	string	Classificação qualitativa	"Positivo"
intensidade	string	Força do sentimento	"Forte"
tweet	string	Trecho do texto analisado	"@ChelseaFC For the first time..."

## 5. Notas Técnicas

### Robustez:

- Verificação explícita da instalação correta dos componentes
- Normalização de scores para evitar valores fora do range

**Extensibilidade:** Pode ser adaptado para outros idiomas trocando o modelo linguístico

**Performance:** Processamento individual por tweet para evitar falhas em cascata